Computer Animation & Motion Capture

Hsin-Chih Lin, Ph.D
Associate Professor
Dept. of Information & Learning Technology
Chief, System & Information Section, Library
National University of Tainan (NUTN)
About Me

Education
- 1986/10 ~ 1990/6 B.E. Electrical Engineering, National Tsing Hua University (NTHU)
- 1990/9 ~ 1997/6 Ph.D. Computer Science, NTHU

Employment History
- 1999/8 ~ 2006/7 Assistant & Associate Professor, Department of Information Management, Chang Jung Christian University (CJCU)
- 2002/8 ~ 2003/7 Chief, Innovation & Incubation Section, Research & Development Office, CJCU
- 2003/8 ~ 2006/7 Director, Extension Program Center, CJCU
About Me (cont.)

- **Research Interests**
  - Visual Information Retrieval (VIR)
  - Computer Graphics & Animation (CG&A)
  - Video Surveillance
  - Management Information Systems (MIS)
  - Digital Archives
國立臺南大學數位典藏聯合目錄
Digital Archives Union Catalogs, National University of Tainan

首頁 / 最新消息 / 聯合目錄 / 計畫說明 / 工作小組

地址：台南市中西區700樹林街二段33號 電話：06-2133111轉294 傳真：06-2149969
Address: 33, Sec. 2, Su-Lin St. Tainan, Taiwan 700 Phone: +886-6-213-3111 # 294 Fax: 06-2149969
Webmaster: feng-ying@mail.nutn.edu.tw Copyright (c) 2007 National University of Tainan all rights reserved
國立臺南大學數位典藏聯合目錄
Digital Archives Union Catalogs, National University of Tainan

日文舊籍珍本

聯合目錄

目前位置 >> 日文舊籍珍本電子書

第 1 / 3 頁

搜索
Outline

- Computer Animation Definition & Milestones
- Computer Animation Production Process
- Motion Capture: Make Avatars Alive
- Content-Based Retrieval for Human Motion Data
- Conclusions & Works in Progress
Computer Animation Definition & Milestones
Computer Graphics & Animation (CG&A)

- Computer graphics (CG) is concerned with all aspects of producing pictures or images using a computer.
- Computer animation is the art of creating moving images via the use of computers. Increasingly it is created by means of 3D CG, though 2D CG is still widely used for low bandwidth and faster real-time needs.
Wireframe

Flat shading

Smooth shading

Texture mapping

Environment mapping

Bump mapping
動畫的定義：孕育生命

- **Animate (vt.)**
  - 賦予生命,使有生命
  - 使活潑,使有活力,使有生氣
  - 激勵,鼓舞
  - 驅動,推動
  - 使栩栩如生地動作
  - 繪製(卡通影片)

- **Animate (a.)**
  - 有生命的;活的
  - 活潑的,有活力的;有生氣的;愉快的

- **Animation (n.)**
  - 生氣,活潑,熱烈[U]
  - 激勵;興奮[U][S1]
  - 活,生(指狀態)[U]
  - 卡通片;卡通片繪製[C][U]
Computer Animation Milestones

- 70′: 剛開始萌芽，著重於技術性的探索，還沒有故事結構或模仿現實的作品
  - (1973 The 1-st SIGGRAPH)
- 80′
- 90′
- 2000~

Utah Teapot (1975)

Beauty & the Beast (1991)

Toy Story (1995)
80'

- **The Advantures of Andre & Willy B. (1984)**
  - John Lasseter in Lucasfilm
  - 第一部真正的電腦動畫短片（有故事元素）

- **Luxo Jr. (1986)**
  - John Lasseter in Pixar
  - 沒有用對白或臉部表情，就成功呈現活生生的幻象
  - 觀眾開始感受到、相信、關心電腦動畫的角色
  - 傳統卡通界與觀眾對電腦動畫為之改觀
  - 第一部獲得奧斯卡提名的電腦動畫短片 (1987)

80’ (cont.)

  - John Lasseter in Pixar
  - 運用燈光、著色和天候特效營造氣氛
  - 電腦動畫開始重視活生生的、有感情的主角（儘管不一定是生物）

- **Technological Threat (1988)**
  - 奧斯卡最佳動畫短片提名
  - 只要動畫基本原則運用得當，電腦動畫與手繪卡通可以成功合成
80' (cont.)

- **Tin Toy (1988)**
  - 第一部獲得奧斯卡獎的 3D 電腦動畫短片

- **The Abyss (1989)**
  - 第一次出現相當逼真的電腦動畫生物
  - 奧斯卡最佳視覺特效

- **Knick Knack (1989)**
  - John Lasseter in Pixar
Pixar’s RenderMan 首次發行 (1989)

- Every movie nominated for Visual Effects Oscar in last 11 years relied on RenderMan
- The first Oscar Award to a software package (March 3, 2001)
90'

  - 奧斯卡最佳視覺特效
  - 把擬真的電腦動畫動作套在真人角色身上

- **Jurassic Park (1993)**
  - 奧斯卡最佳視覺特效
  - 首度演出擬真的電腦動畫生物
  - 讓定格角色 *(stop-motion)* 走入歷史

- **Toy Story (1995)**
  - 第一部電腦動畫劇情長片
  - 奧斯卡提名最佳配樂、最佳主題曲、最佳原著劇本
  - John Lasseter 獲得奧斯卡特殊成就獎
90’ (cont.)

- Woody & Buzz Lightyear 出現在 1996 奧斯卡頒獎典禮
- Dragonheart (1996)
  - 第一部以電腦動畫角色作爲劇情片的主角
- Geri’s Game (1997)
  - 奧斯卡最佳動畫短片 (1997)
  - 把服飾和臉部動畫推進一大步
- Bunny (1998)
  - 奧斯卡最佳動畫短片
  - 燈光氣氛 + 毛髮技術
90’ (cont.)

- **Stuart Little (1999)**
  - 奧斯卡最佳視覺特效
  - 將電腦動畫角色推到劇情片的第一主角

- **For the Birds (2000)**
  - 奧斯卡最佳動畫短片 (2001)

個人動畫製片家逐漸興起
- 軟硬體功能漸大、價格漸低
- 創意與製作足以和大公司競爭
2000~

- Technology is almost completely mature, if you want it, you can make it true.

- Returning to the basics of movie production: Computer animation filmmakers began to focus on story creation, character development, and artistic design, and audiences also walked into theaters to give positive feedback.


- Technology is not the focus, but the strong persuasive story elements. Who has the legend, who wins!
奧斯卡最佳動畫短片提名

生命會有高潮，偶爾也有低潮
當你心情低落就看看四周
你有健康身體，沒痛又沒病
只要換個角度想，世界多美麗

人生在世，起伏不定
只要努力去做，一定會有收穫

2003
奧斯卡最佳動畫短片提名

2005
奧斯卡最佳動畫短片提名
Computer Animation Production Process
最偉大的藝術形式

- 電腦動畫可以運用到其他藝術形式
  - 製圖
  - 繪畫
  - 雕刻
  - 表演
  - 舞蹈
  - 攝影
  - 電影
  - 建築
  - 服飾
  - 佈景設計

- 動畫裡，任何故事都說得出來：沒有任何說故事的媒體能像動畫一樣，提供這麼有彈性的舞台

- 動畫師可以一個人完成這些藝術，esp. 3D 電腦動畫短片
藝術並非大就是美

奧斯卡把短片與其他電影看得同樣重要，影響力一樣大，因為有些想法只要三分鐘就能清楚表達。

好萊塢電影平均用字 17,000，其中 99% 不值一提。

雖然技術進步，但「故事創作」和「角色發展」永遠都是影片的核心。

less is more
動畫短片可以賺錢

- 影展得獎金
- 電視播映權利金
- 影片發行營收
- 在網上直接販賣短片 DVD
- 動畫產業的景氣比其他行業好
- 動畫短片可能發展成爲劇情長片
  - Toy Story, 1995
  - Chicken Run, 2000
  - Ice Age, 2002

* Toy Story, 1995
* Chicken Run, 2000
* Ice Age, 2002
世界三大動畫影展

- ACM SIGGRAPH, USA
  - 2005 Cubic Tragedy: 台科大孫春望教授 + 碩一生全明遠
  - Electronic Theater (電子劇院) 觀眾票選第一名
  - 525 件報名，56 件入選，其中 18 件收錄於 ET 撥映 ($50)，包括：史蒂芬史匹伯「世界大戰」、夢工廠「馬達加斯加」、喬治盧卡斯「星際大戰三部曲」、軍機纏鬥片「機戰未來」

- International Animated Film Festival, Annecy (安錫), France

- Zegreb Film, Croatia
3D 動畫短片製作流程

■ 發展階段 (初期企劃階段)
  □ 故事創作、角色發展、美術設計、分鏡表 (storyboard)

■ 前製階段 (電子元件策劃、創建及合成階段)
  □ 製作企劃 (時程表和預算)、配音、建模、材質與紋理、角色設定、空間構圖、2D 和 3D 預覽片 (animatics)

■ 製作階段 (鏡頭拍攝階段)
  □ 動畫、燈光與著色、視覺特效、合成

■ 後製階段 (完成階段)
  □ 音效與配樂、片名與片單、行銷與發行
製作企劃
Storyboard & 2D/3D Animatics

Sahari

El Arguero

Storyboard

Computer Graphics & Interactive Techniques Lab, NUTN
Animated Short Films: Six Categories

- Elegant Simplicity
- Science Fiction & Fantasy
- Just Plain Fun
- Artsy & Surreal
- On the Darker Side (X)
- On the Lighter Side
Animated Short Films: Six Categories (cont.)

- The Snowman
- run, dragon, run!!
- Grain.S
- "on the sunny side of the street"
- Top Gum
本系大二學生作品
Motion Capture: Make Avatars Alive
Motion Data

- Motion capture (Motion tracking, Mocap): the creation of a 3D representation of a live performance

- Mocap systems
  - Magnetic systems
  - Mechanical systems
  - Optical systems
  - Inertial systems
  - Facial motion capture
  - Performance capture
Magnetic Systems
Mechanical Systems

Gypsy 5
Optical Systems

Camera 1
- 3 Detectors
  - Electronic Shutter
  - 480 Hz
  - 3,000 x 3,000 Optical Resolution
  - 30,000 x 30,000 Sub-Pixel Res
- FPGA Synthetic DSP
  - Sub-Pixel Resolution
  - Ambient Light Subtraction
  - 14-bit position - 8-bit amplitude
  - Horizontal Detector
  - Vertical Detector
  - Diagonal Detector

PHASESPACE
- 12 Cameras
- Active LEDs

HUB Card
- Fits in 5.25" Drive Bay
- Daisy-Chained Cameras
- Dynamic Camera Control

Pentium PC
- Camera Processing Software
- Peak Tracking
- Position, Velocity, Acceleration
- Marker Identification
- Filtering
- Calculate 3D points
- 3D point tracking
- Missing Point Interpolation
- Dynamic Camera Offset
- Display
- Rendering Software
- Editing Software

Computer Graphics & Interactive Techniques Lab, NUTN
Inertial Systems

GypsyGyro 18

Computer Graphics & Interactive Techniques Lab, NUTN
Facial Motion Capture
Performance Capture

- Performance capture captures the body, the hands and facial expression all at the same time. The actor usually interacts with models of the objects in the scene. The recorded motion data can be used to animate different actors.

- In *The Polar Express* (2004), Tom Hanks played six roles.
Advantages & Limitations

- More rapid, sometimes even real time
- Reduce the costs of animation
- Save time and create more natural movements than manual animation
- One actor can play multiple roles within a single film

- No single mocap system is perfect for every possible use
- Limit: anatomically possible motions
The Matrix (1999)

- Bullet time effects
- Digital high-speed camera (12,000fps) + Time-track system
  - Charlie’s Angels (2000)
  - Scary Movie (2000)
  - Shrek (2001)
  - Deuce Bigalow:

Trinity's 360° kick !!

Wow, so many Agent Smith !!

Neo & flying bullets
Content-Based Retrieval for Human Motion Data

Chih-Yi Chiu, Shih-Pin Chao, Ming-Yang Wu, Shi-Nine Yang & Hsin-Chih Lin*
Dept. of CS, National Tsing Hua University
*Dept. of IM, Chang Jung Christian University
Grant & Publications

- **Granted by**
  - MOE Program (Promoting Academic Excellence of Universities) & NSC Project

- **Conference paper**
  - Paper Award

- **Journal paper**
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions

- Content-based retrieval
- Content-based video retrieval
- Investigation of human motion
- Motion data (& motion capture)
- Issues of human motion retrieval

Computer Graphics & Interactive Techniques Lab, NUTN
Content-Based Retrieval

- Today's digital archives & the Web
  - The number of images/videos is growing rapidly
  - We need an efficient technique to effectively find target images/videos

- Content-based retrieval (CBR)
  - Content: color, texture, shape, motion, or the spatial layout in an image (image representation)
  - Users can present the "content of interest" in a query (query expression)
  - Retrieval: find Images/videos that exhibit "perceptual similarity" with the query (matching)
Content-Based Retrieval (cont.)

- Two obstacles in CBR

  □ **Semantic gap**: Users may prefer using high-level textual concepts to interpret an image. However, most of early CBR systems provide low-level numerical features to represent an image.

  □ **Perception subjectivity**: Different users, or even the same user under different circumstances, may interpret an image differently. Moreover, the way users define the similarity between two images may be quite different.
Content-Based Video Retrieval

- Shot detection
  - Abrupt transition (e.g., cuts)
  - Gradual transition (e.g., fades, dissolves & wipes)
- Visual feature extraction (including motion)
- Motion representation
  - MPEG-7's four descriptors: (1) camera motion, (2) motion trajectory, (4) parametric motion, & (4) motion activity
- Indexing
- Matching
Content-Based Video Retrieval

- Single (& rigid) object
- Single motion trajectory
- Relatively short sequence
Investigation of Human Motion

- Computer vision vs computer graphics
- Application fields
  - Visual surveillance
  - Diagnosis & therapy for rehabilitation
  - Athletic training
  - Person identification & verification
  - Animation design & production
  - User interface ...
- We need an efficient approach to search & retrieve human motion data

**Analysis**           **Synthesis**
Motion Data

- Our mocap system is VICON 8 in OES Lab, ITRI, Hsinchu
- We need a human skeleton model
  - H-Anim (Web3D working group on humanoid animation)
  - MPEG-4 body definition parameters (BDPs)
Motion Data (cont.)

Our simplified human skeleton model

- H-Anim 1.0

- 9 skeletal segments & a root orientation

- The default posture is a standing posture with feet pointing to the front, arms on the side and palms facing inward.

- There are 296 Body Animation Parameters (BAPs). When applied to any MPEG-4 compliant generic body, they produce the same animation.

left upper arm
Motion Data (cont.)

\[ D_1: \text{View Direction} \]

\[ D_2 \]

\[ D_3 \]

\[ D_4 \]

\[ D_5 \]

\[ D_6 \]

\[ D_7 \]

\[ D_8 \]
Issues of Human Motion Retrieval

- Our studies are different with conventional content-based video retrieval
  - Single (rigid) object
  - Single motion trajectory
  - Relatively short sequence

- Major problems to be overcome
  - Hierarchical (or articulated) skeletal structure
    => High-dimension feature vectors
  - Multiple motion trajectories
  - Relatively long (consecutive) sequence
Segmented object motion retrieval

Query Clip

Indexing & Matching

Shot A

Shot B

Shot C

Shot D

Similar Clips

Human motion retrieval in a long-length sequence
Issues of Human Motion Retrieval (cont.)

- How to search motion data efficiently?
  - Indexing structures & searching strategies
  - Matching (similarity definition & computation)
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions
Framework Overview

Dynamic Time Warping

Interacting
Outline

- Introduction
- Framework overview
- Indexing
  - Posture representation
- Matching
  - Index map construction
- Experimental results
- Conclusions
Posture Representation

- Local spherical coordinate
- Affine invariance of body transformations
Index Map Construction

- One index map for each body segment
  - Avoid the curse of dimensionality
- Self-organizing map (SOM) clustering
  - Preserve topological properties
  - Improve retrieval accuracy
  - Save matching time
Initial 312 (24x13) cluster centers

Segment-posture distribution

Index map

Left lower arm

Left lower leg

Computer Graphics & Interactive Techniques Lab, NUTN
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions

- Candidate clip searching
- Dynamic time warping
Query Example

Index Map

Start Frame

End Frame

Threshold

Candidate Clip

Computer Graphics & Interactive Techniques Lab, NUTN
Candidate Clip Searching

Advantages

- Reduce time complexity for nearest neighbor searching
- Avoid exhausting search in a long-length sequence
- Hence, matching time is effectively shortened
3 well-known motion recognition techniques

- Template matching: (1) Lower computational cost; (2) more sensitive to length-difference
- Hidden Markov model: (1) Overcome disadv of template matching; (2) training HMM is difficult
- Dynamic time warping: (1) Compensate for length difference & preserve ordering; (2) adv: robust performance & simple computation to match different length patterns
Dynamic Time Warping (cont.)

- **Input**: An example clip & a candidate clip
- **Output**: Similarity between the two clips

We sort the similarity between the example clip & each candidate clip from high to low and then retrieve the top $n$ clips as results.
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions

- Underlying testbed
- Retrieval scenarios
- Retrieval accuracy
- Matching time
Underlying Testbed

- A motion data collection containing 23,130 frames (Tai-Chi Chuan)
  - More than 12 min at 30Hz frame rate
  - A ground truth is established by manually choosing eight groups of motion clips
  - Two clips are relevant if they belong to the same group
- Intel Pentium 4 2.4GHz with 512 MB memory
  - Implemented by Matlab
Retrieval Scenario

Query Input

Retrieval Output

affine-invariance matching
Retrieval Scenario (cont.)

Query Input

Retrieval Output
We implement three indexing methods

- SOM clustering (Proposed)
- Fixed-grid (No clustering)
- $k$-d tree ($k$-means clustering in 18d)

Retrieval accuracy is evaluated by the precision versus recall graph (PR graph)
Retrieval Accuracy (cont.)

- **Precision**

\[
\text{precision} = \frac{\# \{\text{relevant} \cap \text{retrieved}\}}{\# \text{retrieved}}
\]

- **Recall**

\[
\text{recall} = \frac{\# \{\text{relevant} \cap \text{retrieved}\}}{\# \text{relevant}}
\]

- For a given recall value, the precision of the higher curve is better than that of the lower curve.
Retrieval Accuracy (cont.)
**Matching Time**

- **Searching time (candidate clip search)**
- **Similarity computation time (dynamic time warping)**

<table>
<thead>
<tr>
<th></th>
<th>Proposed</th>
<th>Fixed-grid</th>
<th>k-d Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Set Size</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Minimal Window Size</td>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Candidate Clips Searching Time</td>
<td>0.110 sec</td>
<td>0.125 sec</td>
<td>0.531 sec</td>
</tr>
<tr>
<td>Similarity Computation Time</td>
<td>0.187 sec</td>
<td>0.234 sec</td>
<td>0.031 sec</td>
</tr>
<tr>
<td>Total Time</td>
<td>0.297 sec</td>
<td>0.359 sec</td>
<td>0.562 sec</td>
</tr>
</tbody>
</table>
Outline

- Introduction
- Framework overview
- Indexing
- Matching
- Experimental results
- Conclusions
Conclusions

A novel framework for content-based human motion retrieval (a new idea in MIR)

Our studies are different with conventional content-based video retrieval

To improve retrieval performance:
(1) Indexing (2) matching

1.1) Affine invariance representation
1.2) Low-dimension indexing structure
2.1) Heuristic candidate clip searching
2.2) Robust similarity computation

† Single (& rigid) object
† Single motion trajectory
† Relatively short sequence
† Hierarchical (or articulated) skeletal structure
† Multiple motion trajectories
† Relatively long (and consecutive)

82
Conclusions (cont.)

- Better than fixed-grid and $k$-d tree methods in (1) retrieval accuracy (2) matching time
- For CG&A workers, we provide an effective and efficient tool
Conclusions & Works in Progress
Human motion analysis in video
- Give me Wang (face detection/recognition, human detection)
- Event-based video surveillance & retrieval

Human motion synthesis
- Random play & loops
- Interactive play

Interactive Digital Arts / Games
- Human motion as an interface
Conclusions

- 想學電腦動畫，先學說故事
- 故事四大要素：劇情、主角、時間背景、衝突
- 說故事的方法：類型、結構、節奏、製作元件
- 故事元素比製作元素重要
- 故事哪裡來？俯拾皆是！
  - 沙中見世界，花中見菩提
  - 書是良師，請多閱讀；網路不一定是益友，請慎重選擇
  - 隨時尋找生命中的感動

*Computer Graphics & Interactive Techniques Lab, NUTN*
有些事現在不做，一輩子都不會做了
夢想會開始挑戰你的能力、情緒、人際關係，請透過實踐，讓夢想不要質疑你的決心
五月天樂團－阿信

不嗟息過去，更期待未來
不要在乎別人的想法
轉大人，把夢想變大，然後偷偷瘋狂地實現它
但是要看得到
生命展開的寬度和厚度（生命中的感動）
Q&A